ON THE SUBSET SUM PROBLEM OVER FINITE FIELDS

JIYOU LI AND DAQING WAN

Abstract. The subset sum problem over finite fields is a well-known **NP**-complete problem. It arises naturally from decoding generalized Reed-Solomon codes. In this paper, we study the number of solutions of the subset sum problem from a mathematical point of view. In several interesting cases, we obtain explicit or asymptotic formulas for the solution number. As a consequence, we obtain some results on the decoding problem of Reed-Solomon codes.

1. Introduction

Let $\mathbf{F}_q$ be a finite field of characteristic p. Let $D \subseteq \mathbf{F}_q$ be a subset of cardinality $|D| = n > 0$. Let $1 \le m \le k \le n$ be integers. Given m elements $b_1, \cdots, b_m$ in $\mathbf{F}_q$. Let $V_{b,k}$ denote the affine variety in $\mathbf{A}^k$ defined by the following system of equations

$$\sum_{i=1}^{k} X_i = b_1,$$

$$\sum_{1 \le i_1 < i_2 \le k} X_{i_1} X_{i_2} = b_2,$$

$$\cdots,$$

$$\sum_{1 \le i_1 < i_2 < \cdots < i_m \le k} X_{i_1} \cdots X_{i_k} = b_m,$$

$$X_i - X_j \ne 0 \ (i \ne j).$$

A fundamental problem arising from decoding Reed-Solomon codes is to determine for any given $b = (b_1, \cdots, b_m) \in \mathbf{F}_q^m$, if the variety $V_{b,k}$ has an $\mathbf{F}_q$-rational point with all $x_i \in D$, see section 5 for more details. This problem is apparently difficult due to several parameters of different nature involved. The high degree of the variety naturally introduces a substantial algebraic difficulty, but this can at least be overcome in some cases when D is the full field $\mathbf{F}_q$ and m is small, using the Weil bound. The requirement that the x_i's are distinct leads to a significant combinatorial difficulty. From computational point of view, a more substantial difficulty is caused by the flexibility of the subset D of $\mathbf{F}_q$. In fact, even in the case $m = 1$ and so the algebraic difficulty disappear, the problem is known to be **NP**-complete. In this case, the problem is reduced to the well known subset sum problem over $D \subseteq \mathbf{F}_q$, that is, to determine for a given $b \in \mathbf{F}_q$, if there is a non-empty subset $\{x_1, x_2, \cdots, x_k\} \subseteq D$ such that

$$x_1 + x_2 + \cdots + x_k = b. \tag{1.1}$$

This research is partially supported by the NSFC (10331030).

This subset sum problem is known to be **NP**-complete. Given integer $1 \le k \le n$, and $b \in \mathbf{F}_q$, a more precise problem is to determine

$$N(k,b,D) = \#\{\{x_1, x_2, \cdots, x_k\} \subseteq D \mid x_1 + x_2 + \cdots + x_k = b\},$$

the number of k-element subsets of D whose sum is b. The decision version of the above subset sum problem is then to determine if $N(k,b,D) > 0$ for some k, that is, if

$$N(b,D) := \sum_{k=1}^{n} N(k,b,D) > 0.$$

In this paper, we study the approximation version of the above subset sum problem for each k from a mathematical point of view, that is, we try to approximate the solution number $N(k,b,D)$. Intuitively, the problem is easier if D is close to be the full field $\mathbf{F}_q$, i.e., when $q-n$ is small. Indeed, we obtain an asymptotic formula for $N(k,b,D)$ when $q-n$ is small. Heuristically, $N(k,b,D)$ should be approximately $\frac{1}{q}\binom{n}{k}$. The question is about the error term. We have

Theorem 1.1. *Let $p < q$, that is, $\mathbf{F}_q$ is not a prime field. Let $D \subseteq \mathbf{F}_q$ be a subset of cardinality n. For any $1 \le k \le n \le q-2$, any $b \in \mathbf{F}_q$, we have the inequality*

$$\left| N(k,b,D) - \frac{1}{q}\binom{n}{k} \right| \le \frac{q-p}{q} \binom{k+q-n-2}{q-n-2} \binom{q/p-1}{\lfloor k/p \rfloor}.$$

Furthermore, let $D = \mathbf{F}_q \backslash \{a_1, \cdots, a_{q-n}\}$ with $a_1 = 0$, and if $b, a_2, \cdots, a_{q-n}$ are linearly independent over $\mathbf{F}_p$, then we have the improved estimate

$$\left| N(k,b,D) - \frac{1}{q}\binom{n}{k} \right| \le \max_{0 \le j \le k} \frac{p}{q} \cdot \binom{k+q-n-2-j}{q-n-2} \binom{q/p-1}{\lfloor j/p \rfloor}.$$

When $q = p$, that is, $\mathbf{F}_q$ is a prime field, we have

$$\left| N(k,b,D) - \frac{1}{q}\binom{n}{k} + \frac{(-1)^k}{q} \binom{k+q-n-1}{q-n-1} \right| \le \binom{k+q-n-2}{q-n-2}.$$

Theorem 1.1 assumes that $n \le q-2$. In the remaining case $n \ge q-2$, that is, $n \in \{q-2, q-1, q\}$, the situation is nicer and we obtain explicit formulas for $N(k,b,D)$. Here we first state the results for $q-n \le 1$ and thus we can take $D = \mathbf{F}_q$ or $\mathbf{F}_q^*$.

Theorem 1.2. *Define $v(b) = -1$ if $b \ne 0$, and $v(b) = q-1$ if $b = 0$. Then*

$$N(k,b,\mathbf{F}_q^*) = \frac{1}{q}\binom{q-1}{k} + (-1)^{k+\lfloor k/p \rfloor} \frac{v(b)}{q} \binom{q/p-1}{\lfloor k/p \rfloor}.$$

If $p \nmid k$, then

$$N(k,b,\mathbf{F}_q) = \frac{1}{q}\binom{q}{k}.$$

If $p \mid k$, then

$$N(k,b,\mathbf{F}_q) = \frac{1}{q}\binom{q}{k} + (-1)^{k+\frac{k}{p}} \frac{v(b)}{q} \binom{q/p}{k/p}.$$

When $q-n = 2$, note that we can always take $D = \mathbf{F}_q \backslash \{0,1\}$.

Theorem 1.3. *Let $q > 2$. Then we have*

$$N(k, b, \mathbf{F}_q \backslash \{0, 1\}) = \frac{1}{q}\binom{q-2}{k} + \frac{1}{q}(-1)^k R_k^2 - (-1)^k S(k, k-b),$$

where $R_k^2, S(k, b)$ are defined as in (3.2) and (3.3).

This paper is organized as follows: We first prove Theorem 1.2 and Theorem 1.3 in Section 2 and Section 3 respectively. Then we prove Theorem 1.1 in Section 4. Applications to coding theory are given in Section 5.

Notations. For $x \in \mathbb{R}$, let $(x)_0 = 1$ and $(x)_k = x(x-1)\cdots(x-k+1)$ for $k \in \mathbb{Z}^+ = \{1, 2, 3, \cdots\}$. For $k \in \mathbb{N} = \{0, 1, 2, \cdots\}$ define the binomial coefficient $\binom{x}{k} = \frac{(x)_k}{k!}$. For a real number a we denote $\lfloor a \rfloor$ to be the largest integer not greater than a.

2. Proof of Theorem 1.2

When D equals $q-1$, it suffices to consider $N(k, b, \mathbf{F}_q^*)$ by a simple linear substitution. Let $M(k, b, D)$ denote the number of ordered tuples $(x_1, x_2, \cdots, x_k)$ satisfying equation (1.1). Then

$$M(k, b, D) = k! N(k, b, D)$$

is the number of solutions of the equation

$$x_1 + \cdots + x_k = b, x_i \in D, x_i \neq x_j \ (i \neq j). \tag{2.1}$$

It suffices to determine $M(k, b, D)$. We use a pure combinatorial method to find recursive relations among the values of $M(k, b, \mathbf{F}_q)$ and $M(k, b, \mathbf{F}_q^*)$.

Lemma 2.1. *For $b \neq 0$ and D being $\mathbf{F}_q$ or $\mathbf{F}_q^*$, we have $M(k, b, D) = M(k, 1, D)$.*

Proof. There is a one to one map sending the solution $\{x_1, x_2, \cdots, x_k\}$ of (2.1) to the solution $\{x_1 b^{-1}, x_2 b^{-1}, \cdots, x_k b^{-1}\}$ of (2.1) with $b = 1$. □

Lemma 2.2.

$$M(k, 1, \mathbf{F}_q) = M(k, 1, \mathbf{F}_q^*) + kM(k-1, 1, \mathbf{F}_q^*), \tag{2.2}$$

$$M(k, 0, \mathbf{F}_q) = M(k, 0, \mathbf{F}_q^*) + kM(k-1, 0, \mathbf{F}_q^*), \tag{2.3}$$

$$(q)_k = (q-1)M(k, 1, \mathbf{F}_q) + M(k, 0, \mathbf{F}_q), \tag{2.4}$$

$$(q-1)_k = (q-1)M(k, 1, \mathbf{F}_q^*) + M(k, 0, \mathbf{F}_q^*). \tag{2.5}$$

Proof. Fix an element $c \in \mathbf{F}_q$. The solutions of (2.1) in $\mathbf{F}_q$ can be divided into two classes depending on whether c occurs. By a linear substitution, the number of solutions of (2.1) in $\mathbf{F}_q$ not including c equals $M(k, b-ck, \mathbf{F}_q^*)$. And the number of solutions of (2.1) in $\mathbf{F}_q$ including c equals $kM(k-1, b-ck, \mathbf{F}_q^*)$. Hence we have

$$M(k, b, \mathbf{F}_q) = M(k, b-ck, \mathbf{F}_q^*) + kM(k-1, b-ck, \mathbf{F}_q^*). \tag{2.6}$$

Then (2.2) follows by choosing $b = 1, c = 0$. Similarly, (2.3) follows by choosing $b = 0, c = 0$. Note that $(q)_k$ is the number of k-permutations of $\mathbf{F}_q$, and $(q-1)_k$ is the number of k-permutations of $\mathbf{F}_q^*$. Thus, both (2.4) and (2.5) follows. □

The next step is to find more relations between $M(k, b, \mathbf{F}_q)$ and $M(k, b, \mathbf{F}_q^*)$.

Lemma 2.3. *If $p \nmid k$, we have $M(k,b,\mathbf{F}_q) = M(k,0,\mathbf{F}_q)$ for all $b \in \mathbf{F}_q$ and hence*

$$M(k,b,\mathbf{F}_q) = \frac{1}{q}(q)_k.$$

If $p \mid k$, we have $M(k,b,\mathbf{F}_q) = qM(k-1,b,\mathbf{F}_q^*)$ for all $b \in \mathbf{F}_q$.

Proof. **Case 1:** Since $p \nmid k$, we can take $c = k^{-1}b$ in (2.6) and get the relation

$$M(k,b,\mathbf{F}_q) = M(k,0,\mathbf{F}_q^*) + kM(k-1,0,\mathbf{F}_q^*).$$

The right side is just $M(k,0,\mathbf{F}_q)$ by (2.3).

Case 2 : In this case, $p \mid k$. Then $M(k,b,\mathbf{F}_q)$ equals the number of ordered solutions of the following system of equations:

$$\begin{cases} x_1 + x_2 + \cdots + x_k = b, \\ x_1 - x_2 = y_2, \\ \cdots\cdots \\ x_1 - x_k = y_k, \\ y_i \in \mathbf{F}_q^*, \quad y_i \neq y_j,\ 2 \le i < j \le k. \end{cases}$$

Regarding $x_1, x_2, \cdots, x_k$ as variables it is easy to check that the p-rank (the rank of a matrix over the prime field $\mathbf{F}_p$) of the coefficient matrix of the above system of equations equals $k-1$. The system has solutions if and only if $\sum_{i=2}^{k} y_i = -b$ and $y_i \in \mathbf{F}_q^*$ being distinct. Furthermore, since the p-rank of the above system is $k-1$, when $y_2, y_3, \cdots, y_k$ and x_1 are given then $x_2, x_3, \cdots, x_k$ will be uniquely determined. This means the number of the solutions of above linear system of equations equals to q times the number of ordered solutions of the following equation:

$$\begin{cases} y_2 + y_3 + \cdots + y_k = -b, \\ y_i \in \mathbf{F}_q^*, \quad y_i \neq y_j,\ 2 \le i < j \le k. \end{cases}$$

This number of solutions of the above equation is just $M(k-1,b,\mathbf{F}_q^*)$ and hence $M(k,b,\mathbf{F}_q) = qM(k-1,b,\mathbf{F}_q^*)$. □

We have obtained several relations from Lemma 2.2 and Lemma 2.3. To determine $M(k,b,\mathbf{F}_q)$, it is now sufficient to know $M(k,0,\mathbf{F}_q^*)$. Define for $k > 0$,

$$d_k = M(k,1,\mathbf{F}_q^*) - M(k,0,\mathbf{F}_q^*).$$

Then by (2.5) we have

$$qM(k,0,\mathbf{F}_q^*) = (q-1)_k - (q-1)d_k. \tag{2.7}$$

Heuristically, $M(k,0,\mathbf{F}_q^*)$ should be approximately $\frac{1}{q}(q-1)_k$. To obtain the explicit value of $M(k,0,\mathbf{F}_q^*)$, we only need to know d_k. For convenience we set $d_0 = -1$.

Lemma 2.4. *If d_k is defined as above, then*

$$d_k = \begin{cases} -1, & k = 0; \\ 1, & k = 1; \\ -kd_{k-1}, & p \nmid k,\ 2 \le k \le q-1\ ; \\ (q-k)d_{k-1}, & p \mid k,\ 2 \le k \le q-1\ . \end{cases}$$

Proof. One checks that $d_1 = M(1,1,\mathbf{F}_q^*) - M(1,0,\mathbf{F}_q^*) = 1 - 0 = 1$. When $p \nmid k$, by Lemma 2.3 we have $M(k,1,\mathbf{F}_q) = M(k,0,\mathbf{F}_q)$. This together with Lemma 2.2 implies

$$M(k,1,\mathbf{F}_q^*) - M(k,0,\mathbf{F}_q^*) = k(M(k-1,0,\mathbf{F}_q^*) - M(k-1,1,\mathbf{F}_q^*)).$$

Namely, $d_k = -kd_{k-1}$. When $p \mid k$, using Lemma 2.3 we have

$$M(k,1,\mathbf{F}_q) - M(k,0,\mathbf{F}_q) = q(M(k-1,1,\mathbf{F}_q^*) - M(k-1,0,\mathbf{F}_q^*)) = qd_{k-1}.$$

By Lemma 2.2, the left side is $d_k + kd_{k-1}$. Thus, $d_k = (q-k)d_{k-1}$. □

Corollary 2.5.

$$d_k = -(-1)^{k+\lfloor k/p \rfloor} k! \binom{q/p-1}{\lfloor k/p \rfloor}.$$

Proof. One checks $d_0 = -1$ and $d_1 = 1$ are consistent with the above formula for $k \leq 1$. Let $k \geq 2$ and write $k = np + m$ with $0 \leq m < p$. By Lemma 2.4,

$$\begin{aligned} \frac{d_k}{k!} &= (-1)^{n(p-1)+m+1} \prod_{i=1}^{n} \frac{(q-ip)}{ip} \\ &= (-1)^{n(p-1)+m+1} \frac{\prod_{i=1}^{n} (q/p-i)}{n!} \\ &= -(-1)^{k+n} \binom{q/p-1}{n}. \end{aligned}$$

It is easy to check that if $q = p$, then we have $d_k = (-1)^{k-1}k!$, which is consistent with the definition $(0)_0 = 1$. □

Proof of Theorem 1.2 Let $M(k,b,D)$ be the number of solutions of (2.1). Note that $M(k,b,D) = k!N(k,b,D)$ and $d_k = -(-1)^{k+\lfloor k/p \rfloor} k! \binom{q/p-1}{\lfloor k/p \rfloor}$. Thus it is sufficient to prove

$$\begin{aligned} M(k,b,\mathbf{F}_q^*) &= \frac{(q-1)_k - v(b)d_k}{q}; \\ M(k,b,\mathbf{F}_q) &= \frac{(q)_k - v(b)(d_k + kd_{k-1})}{q}. \end{aligned}$$

If $b = 0$, by (2.7), we obtain

$$qM(k,0,\mathbf{F}_q^*) = (q-1)_k - (q-1)d_k.$$

If $b \neq 0$, then

$$qM(k,b,\mathbf{F}_q^*) = qM(k,1,\mathbf{F}_q^*) = qd_k + qM(k,0,\mathbf{F}_q^*) = (q-1)_k + d_k.$$

The formula for $M(k,b,\mathbf{F}_q^*)$ holds.

If $p \nmid k$, then $d_k + kd_{k-1} = 0$ and the formula for $M(k,b,\mathbf{F}_q)$ holds by Lemma 2.3.

If $p \mid k$, then $d_k + kd_{k-1} = qd_{k-1}$. By Lemma 2.3 and the above formula for $M(k,b,\mathbf{F}_q^*)$, we deduce

$$M(k,b,\mathbf{F}_q) = qM(k-1,b,\mathbf{F}_q^*) = (q-1)_{k-1} - v(b)d_{k-1}.$$

The formula for $M(k,b,\mathbf{F}_q)$ holds. The proof is complete.

Now we turn to deciding when the solution number $N(k,b,\mathbf{F}_q^*) > 0$. A sequence $\{a_0, a_1, \cdots, a_n\}$ is **unimodal** if there exits index k with $0 \leq k \leq n$ such that

$$a_0 \leq a_1 \leq \cdots a_{k-1} \leq a_k \geq a_{k+1} \cdots \geq a_n.$$

The sequence $\{a_0, a_1, \cdots, a_n\}$ is called symmetric if $a_i = a_{n-i}$ for $0 \leq i < n$.

Corollary 2.6. *For any $b \in \mathbf{F}_q$, both the sequence $N(k, b, \mathbf{F}_q)$ $(1 \le k \le q)$ and the sequence $N(k, b, \mathbf{F}_q^*)$ $(1 \le k \le q-1)$ are unimodal and symmetric.*

Proof. The symmetric part can be verified using Theorem 1.1. A simpler way is to use the relation

$$\sum_{a \in \mathbf{F}_q} a = \sum_{a \in \mathbf{F}_q^*} a = 0.$$

To prove the unimodal property for $N(k, b, \mathbf{F}_q^*)$, by the symmetry it is sufficient to consider the case $k \le \frac{q-1}{2}$. Then, by Theorem 1.1, we deduce

$$\begin{aligned} & q\left(N(k,0,\mathbf{F}_q^*) - N(k-1,0,\mathbf{F}_q^*)\right) \\ & \ge \binom{q-1}{k} - \binom{q-1}{k-1} - (q-1)\left(\binom{q/p-1}{\lfloor k/p \rfloor} - \binom{q/p-1}{\lfloor (k-1)/p \rfloor}\right). \end{aligned}$$

If $p \nmid k$, then $\lfloor k/p \rfloor = \lfloor (k-1)/p \rfloor$ and the right side is clearly positive. If $p \mid k$, then

$$\begin{aligned} & q\left(N(k,0,\mathbf{F}_q^*) - N(k-1,0,\mathbf{F}_q^*)\right) \\ & \ge \frac{q-2k}{k}\binom{q-1}{k-1} - (q-1)\frac{q/p-2k/p}{k/p}\binom{q/p-1}{k/p-1} \\ & = \frac{q-2k}{k}\left(\binom{q-1}{k-1} - (q-1)\binom{q/p-1}{k/p-1}\right). \end{aligned} \tag{2.8}$$

When $p = 2$ and $k = 2, 4$, or $q \le 9$, it is easy to checks that $\binom{q-1}{k-1} \ge (q-1)\binom{q/p-1}{k/p-1}$. Otherwise by the Vandermonde's convolution

$$\binom{q-1}{k-1} = \sum_{i=0}^{q/p-1} \binom{q/p-1}{i}\binom{q-q/p}{k-1-i},$$

it suffices to prove

$$\binom{q-q/p}{k-k/p} \ge q-1.$$

This inequality follows by noting that

$$\binom{q-q/p}{k-k/p} \ge \binom{q/2}{2}$$

and $q > 9$. Thus $N(k, 0, \mathbf{F}_q^*)$ is unimodal. The proof for the unimodality of $N(k, b, \mathbf{F}_q)$ is similar. This completes the proof.

□

Corollary 2.7. *Let $|D| = q-1 > 4$. If p is an odd prime then for $1 < k < q-2$ the equation* (1.1) *always has a solution. If $p = 2$, then for $2 < k < q-3$ the equation* (1.1) *always has a solution.*

Proof. For any $a \in \mathbf{F}_q$ we have $N(k, b, \mathbf{F}_q \backslash \{a\}) = N(k, b - ka, \mathbf{F}_q^*)$. Thus it is sufficient to consider $N(k, 1, \mathbf{F}_q^*)$ and $N(k, 0, \mathbf{F}_q^*)$ by Lemma 2.1. When p is odd and $k = 2$, we have $N(2, 0, \mathbf{F}_q^*) = \frac{1}{q}\left(\binom{q-1}{2} + (q-1)\right) = \frac{q-1}{2} > 0$, and $N(2, 1, \mathbf{F}_q^*) = \frac{1}{q}\left(\binom{q-1}{2} - 1\right) = \frac{q-3}{2} > 0$ from Theorem 1.2. Then, by the unimodality of $N(k, 1, \mathbf{F}_q^*)$ and $N(k, 0, \mathbf{F}_q^*)$, for $1 < k < q-2$, $N(k, b, \mathbf{F}_q \backslash \{a\})$ must be positive.

Similarly, when $p = 2$ and $k = 3$ we have $N(3, 0, \mathbf{F}_q^*) = \frac{1}{q}\left(\binom{q-1}{3} + (q-1)(\frac{q}{2} - 1)\right) = \frac{(q-1)(q-2)}{6} > 0$ and $N(3, 1, \mathbf{F}_q^*) = \frac{1}{q}\left(\binom{q-1}{3} - (\frac{q}{2} - 1)\right) = \frac{(q-2)(q-4)}{6} > 0$. By the unimodality and symmetry we complete the proof. □

Corollary 2.8. *Let $D = \mathbf{F}_q$. If p is an odd prime then the equation (1.1) always has a solution if and only if $0 < k < q$. If $p = 2$, then for $2 < k < q-2$ the equation (1.1) always has a solution.*

Proof. It is straightforward from Corollary 2.7 and Theorem 1.1. □

3. Proof of Theorem 1.3

Before our proof of Theorem 1.3, we first give several lemmas, which give some basic formulas for the summands of sign-alternating binomial coefficients.

Lemma 3.1. *Let k, m be integers. Then we have*

$$\sum_{k\le m}(-1)^k\binom{r}{k} = (-1)^m\binom{r-1}{m}.$$

Proof. It follows by comparing the coefficients of x^m in both sides of $(1-x)^{-1}(1-x)^r = (1-x)^{r-1}$. □

Lemma 3.2. *Let $< k >_p$ be the least non-negative residue of k modulo p. For any positive integers a, k, we have*

$$\sum_{j=0}^{k} -(-1)^{\lfloor j/p\rfloor}\binom{a}{\lfloor j/p\rfloor} = -p(-1)^{\lfloor k/p\rfloor}\binom{a-1}{\lfloor k/p\rfloor}+(p-1-< k >_p)(-1)^{\lfloor k/p\rfloor}\binom{a}{\lfloor k/p\rfloor},$$

and thus

$$\sum_{j=0}^{k} -(-1)^{\lfloor j/p\rfloor}\binom{a}{\lfloor j/p\rfloor} \le p\binom{a}{\lfloor k/p\rfloor}. \tag{3.1}$$

Proof. Let $j = n_j p + m_j$ with $0 \le m_j < p$. Applying Lemma 3.1 we have

$$\begin{aligned}
&\sum_{j=0}^{k} -(-1)^{\lfloor \frac{j}{p}\rfloor}\binom{a}{\lfloor j/p\rfloor}\\
&= -p\sum_{n_j=0}^{n_k}(-1)^{n_j}\binom{a}{n_j} + (p-1-< k >_p)(-1)^{n_k}\binom{a}{n_k}\\
&= -p(-1)^{\lfloor k/p\rfloor}\binom{a-1}{\lfloor k/p\rfloor} + (p-1-< k >_p)(-1)^{\lfloor k/p\rfloor}\binom{a}{\lfloor k/p\rfloor}.
\end{aligned}$$

The inequality (3.1) follows by noting the alternating signs before the two binomial coefficients. □

Lemma 3.3. *Let $R_k^1 = (-1)^k\frac{d_k}{k!} = -(-1)^{\lfloor k/p\rfloor}\binom{q/p-1}{\lfloor k/p\rfloor}$. Let $< k >_p$ denote the least non-negative residue of k modulo p. Define $R_k^2 = \sum_{j=0}^{k} R_j^1$. Then we have*

$$R_k^2 = -p(-1)^{\lfloor k/p\rfloor}\binom{q/p-2}{\lfloor k/p\rfloor} + (p-1-< k >_p)(-1)^{\lfloor k/p\rfloor}\binom{q/p-1}{\lfloor k/p\rfloor}. \tag{3.2}$$

Moreover, let $b \in \mathbf{F}_p$. Define $\delta_{b,k} = 1$ if $< b >_p$ is greater than $< k >_p$ and $\delta_{b,k} = 0$ otherwise. Then we have

$$S(k,b) := \sum_{\substack{0\le i\le k\\ i\equiv b(\bmod p)}} R_i^1 = -(-1)^{\lfloor k/p\rfloor}\binom{q/p-2}{\lfloor k/p\rfloor} + \delta_{b,k}(-1)^{\lfloor k/p\rfloor}\binom{q/p-1}{\lfloor k/p\rfloor}. \tag{3.3}$$

Proof. Note that (3.2) is direct from Lemma 3.2 by setting $a = q/p - 1$. Since it is similar to that of Lemma 3.2, we omit the proof of (3.3). □

We extend the equation (3.3) by defining $S(k, b) = 0$ for $b \notin \mathbf{F}_p$ and any integer k. Note that $S(k, b) \leq \binom{q/p-2}{\lfloor k/p \rfloor}$. In the following theorem, we give the accurate formula for $N(k, b, D)$ when $D = \mathbf{F}_q \backslash \{a_1, a_2\}$ and first note that we can always assume $a_1 = 0$ and $a_2 = 1$ by a linear substitution.

Proof of Theorem 1.3 Using the simple inclusion-exclusion sieving method by considering whether a_2 appears in the solution of equation (1.1) we have

$$\begin{aligned}
&N(k, b, \mathbf{F}_q \backslash \{a_1, a_2\}) \\
&= N(k, b, \mathbf{F}_q \backslash \{a_1\}) - N(k-1, b-a_2, \mathbf{F}_q \backslash \{a_1, a_2\}) \\
&= N(k, b, \mathbf{F}_q \backslash \{a_1\}) - (N(k-1, b-a_2, \mathbf{F}_q \backslash \{a_1\}) \\
&\quad - N(k-2, b-2a_2, \mathbf{F}_q \backslash \{a_1, a_2\})) \\
&\quad \cdots\cdots \\
&= \sum_{i=0}^{k-1} (-1)^i N(k-i, b-ia_2, \mathbf{F}_q \backslash \{a_1\}) \\
&\quad + (-1)^k N(0, b-ka_2, \mathbf{F}_q \backslash \{a_1, a_2\}).
\end{aligned}$$

One checks that the above equation holds if we define $N(0, b, D)$ to be 1 if and only if $b = 0$ for a nonempty set D. Noting that $a_1 = 0$ we have

$$N(k, b, \mathbf{F}_q \backslash \{a_1, a_2\}) = \sum_{i=0}^{k} (-1)^i N(k-i, b-ia_2, \mathbf{F}_q^*).$$

From Theorem 1.1 we have the following formula

$$N(k, b, \mathbf{F}_q^*) = \frac{1}{q}\binom{q-1}{k} - \frac{1}{q}(-1)^k v(b) R_k^1,$$

where $R_k^1 = -(-1)^{\lfloor k/p \rfloor}\binom{q/p-1}{\lfloor k/p \rfloor}$, $v(b) = -1$ if $b \neq 0$ and $v(b) = q-1$ if $b = 0$. Thus

$$\begin{aligned}
&N(k, b, \mathbf{F}_q \backslash \{a_1, a_2\}) \\
&= \sum_{i=0}^{k} (-1)^i \left(\frac{1}{q}\binom{q-1}{k-i} - \frac{1}{q}(-1)^{k-i} v(b-ia_2) R_{k-i}^1 \right). \\
&= \frac{1}{q}\left((-1)^k \sum_{k-i=0}^{k} (-1)^{k-i}\binom{q-1}{k-i} - (-1)^k \sum_{k-i=0}^{k} v(b-ia_2) R_{k-i}^1 \right) \\
&= \frac{1}{q}\left((-1)^k \sum_{j=0}^{k} (-1)^j \binom{q-1}{j} - (-1)^k \sum_{j=0}^{k} v(b-ka_2+ja_2) R_j^1 \right) \\
&= \frac{1}{q}\left(\binom{q-2}{k} - (-1)^k \sum_{j=0}^{k} v(b-ka_2+ja_2) R_j^1 \right).
\end{aligned}$$

The last equality follows from Lemma 3.1. Noting that $a_2 = 1$, and by the definition of $v(b)$ we have

$$N(k, b, \mathbf{F}_q \backslash \{a_1, a_2\})$$

$$= \frac{1}{q}\binom{q-2}{k} - \frac{1}{q}(-1)^k \sum_{j=0}^{k} v(b-k+j) R_j^1 \tag{3.4}$$

$$= \frac{1}{q}\binom{q-2}{k} - \frac{1}{q}(-1)^k \sum_{j=0}^{k} (-1)\cdot R_j^1 - \frac{1}{q}(-1)^k \sum_{\substack{0\le j\le k \\ b-k+j=0}} q \cdot R_j^1$$

$$= \frac{1}{q}\binom{q-2}{k} + \frac{1}{q}(-1)^k R_k^2 - (-1)^k \cdot S(k, k-b). \tag{3.5}$$

The proof is complete.

Combining (3.4), (3.2) and (3.3) we obtain the following simple solution number formula compared with those stated in Theorem 1.2 and Theorem 1.3.

Corollary 3.4. *If $< k >_p = p-1$ and $b \in \mathbf{F}_p$, then we have*

$$N(k, b, \mathbf{F}_q \backslash \{0,1\}) = \frac{1}{q}\binom{q-2}{k} + (-1)^{k+\lfloor k/p \rfloor} \frac{q-p}{q} \binom{q/p-2}{\lfloor k/p \rfloor}.$$

This shows that the estimate in Theorem 1.1 is nearly sharp for $q - n = 2$.

4. Proof of Theorem 1.1

Let $D = \mathbf{F}_q \backslash \{a_1, a_2, \cdots a_c\}$, where $a_1, a_2, \cdots a_c$ are distinct elements in $\mathbf{F}_q$. In this section, based on the explicit formula of $N(k, b, D)$ for $c = 2$ given in Theorem 1.3, we first obtain a general formula for $c > 2$. Then we give the proof of Theorem 1.1. The solution number $N(k, b, \mathbf{F}_q \backslash \{a_1, a_2, \cdots a_c\})$ is closely related to the $\mathbf{F}_p$-linear relations among the set $\{a_1, \cdots a_c\}$ which we will see in Lemma 4.2. For the purpose of Theorem 1.1's proof and further investigations on the solution number $N(k, b, D)$, we first state the following lemma.

Lemma 4.1. *Let $R_k^1 = -(-1)^{\lfloor k/p \rfloor}\binom{q/p-1}{\lfloor k/p \rfloor}$. For $c > 1$ if we define recursively that $R_k^c = \sum_{j=0}^{k} R_j^{c-1}$, then we have*

$$R_k^c = -\sum_{j=0}^{k} (-1)^{\lfloor j/p \rfloor} \binom{k+c-2-j}{c-2} \binom{q/p-1}{\lfloor j/p \rfloor}. \tag{4.1}$$

Proof. When $c = 2$, this formula is just the definition of R_k^2. Assume it is true for some $c \ge 2$, then we have

$$\begin{aligned}
R_k^{c+1} &= \sum_{i=0}^{k} R_i^c \\
&= \sum_{i=0}^{k} (-1) \cdot \sum_{j=0}^{i} (-1)^{\lfloor j/p \rfloor} \binom{i+c-2-j}{c-2} \binom{q/p-1}{\lfloor j/p \rfloor} \\
&= -\sum_{j=0}^{k} \sum_{i=j}^{k} (-1)^{\lfloor j/p \rfloor} \binom{i+c-2-j}{c-2} \binom{q/p-1}{\lfloor j/p \rfloor} \\
&= -\sum_{j=0}^{k} (-1)^{\lfloor j/p \rfloor} \sum_{i=j}^{k} \binom{i+c-2-j}{c-2} \binom{q/p-1}{\lfloor j/p \rfloor}
\end{aligned}$$

$$= -\sum_{j=0}^{k}(-1)^{\lfloor j/p\rfloor}\binom{k+c-1-j}{c-1}\binom{q/p-1}{\lfloor j/p\rfloor}.$$

The last equality follows from the following simple binomial coefficient identity

$$\sum_{j\le k}\binom{j+n}{n} = \binom{k+n+1}{n+1}.$$

□

It is easy to check that when $k > \frac{q-c}{2}$, we have

$$N(k,b,D) = N(q-c-k, -b-\sum_{i=1}^{c} a_i, D),$$

where $D = \mathbf{F}_q\backslash\{a_1, a_2, \cdots, a_c\}$. Thus we may always assume that $k \le \frac{q-c}{2}$. In the following lemma, for convenience we state two different types of formulas.

Lemma 4.2. *Let $D = \mathbf{F}_q\backslash\{a_1, a_2, \cdots, a_c\}$ and $c \ge 3$, where $a_1 = 0, a_2 = 1, a_3, \cdots, a_c$ are distinct elements in the finite field $\mathbf{F}_q$ of characteristic p. Define the integer valued function $v(b) = -1$ if $b \ne 0$ and $v(b) = q-1$ if $b = 0$. Then for any $b \in \mathbf{F}_q$, we have the formulas*

$$\begin{aligned}
&N(k,b,D) - \frac{1}{q}\binom{q-c}{k}\\
&= -\frac{1}{q}(-1)^k \cdot \sum_{i_1=0}^{k}\sum_{i_2=0}^{k-i_1}\cdots\sum_{i_{c-1}=0}^{k-i_1-\cdots-i_{c-2}} v(b - i_1a_c - \cdots - (k-\sum_{j=1}^{c-1} i_j)a_2)R_j^1 \qquad (4.2)\\
&= \frac{1}{q}(-1)^k R_k^c - (-1)^k \cdot \sum_{i_1=0}^{k}\sum_{i_2=0}^{k-i_1}\cdots\sum_{i_{c-2}=0}^{k-i_1-\cdots-i_{c-3}}\\
&\qquad S(k-\sum_{j=1}^{c-2} i_j, k-\sum_{j=1}^{c-2} i_j - b + \sum_{j=1}^{c-2} i_j a_{c+1-j}), \qquad (4.3)
\end{aligned}$$

where R_k^c is defined by (3.2), and $S(k,b)$ is defined by (3.3). Moreover, if $a_1 = 0$, and $b, a_2, \cdots, a_c$ are linear independent over $\mathbf{F}_p$, then we have

$$N(k,b,D) = \frac{1}{q}\binom{q-c}{k} + \frac{1}{q}(-1)^k R_k^c. \qquad (4.4)$$

Proof. Using the simple inclusion-exclusion sieving method we have

$$\begin{aligned}
&N(k,b,\mathbf{F}_q\backslash\{a_1,a_2,\cdots,a_c\})\\
&= N(k,b,\mathbf{F}_q\backslash\{a_1,a_2,\cdots,a_{c-1}\})\\
&\quad -N(k-1,b-a_c,\mathbf{F}_q\backslash\{a_1,a_2,\cdots,a_{c-1}\})\\
&= \cdots\cdots\\
&= \sum_{i=0}^{k}(-1)^i N(k-i, b-ia_c, \mathbf{F}_q\backslash\{a_1, a_2\cdots, a_{c-1}\}).
\end{aligned}$$

When $c = 3$, noting that $a_2 = 1$, (3.5) implies that

$$N(k,b,\mathbf{F}_q\backslash\{a_1,a_2,a_3\})$$

$$= \sum_{i=0}^{k}(-1)^i\left(\frac{1}{q}\binom{q-2}{k-i}+\frac{1}{q}(-1)^{k-i}R_{k-i}^2-(-1)^{k-i}S(k-i,k-i-(b-ia_3))\right)$$

$$= \frac{1}{q}\binom{q-3}{k}+\frac{1}{q}(-1)^kR_k^3-(-1)^k\sum_{i=0}^{k}S(k-i,k-i-b+ia_3).$$

By induction, (4.3) follows for $c \geq 3$. Similarly, (4.2) follows from (3.4).

If $b, a_2 = 1, a_3 \cdots, a_c$ are linear independent over $\mathbf{F}_p$, then first note that $b \notin \mathbf{F}_p$. Thus, when $c = 2$, by its extended definition we have $S(k, k-b) = 0$ for any integer k. When $c > 2$, since $b, a_2 = 1, a_3 \cdots, a_c$ are independent, we know that $k - \sum_{j=1}^{c-2} i_j - b + \sum_{j=1}^{c-2} i_j a_{c+1-j} \notin \mathbf{F}_p$ for any index tuple $(i_1, i_2, \cdots, i_{c-2})$ in the summation of (4.3). Thus this summation always vanishes for any c and the proof is complete. □

Now we have obtained the two formulas of the solution number $N(k, b, D)$. It suffices to evaluate R_k^c and the summation in (4.3), which is denoted by S_k^c . Unfortunately, S_k^c is extremely complicated when c is large. The **NP**-hardness of the subset sum problem indicates the hardness of precisely evaluating it. In the following lemmas we first deduce a simple bounds for R_k^c and S_k^c.

Lemma 4.3. *Let $p < q$. Let*

$$S_k^c = \sum_{i_1=0}^{k}\sum_{i_2=0}^{k-i_1}\cdots\sum_{i_{c-2}=0}^{k-i_1-\cdots-i_{c-3}} S(k-\sum_{j=1}^{c-2}i_j, k-\sum_{j=1}^{c-2}i_j-b+\sum_{j=1}^{c-2}i_ja_{c+1-j}).$$

Then we have

$$qS_k^c - R_k^c \leq (q-p)\binom{k+c-2}{c-2}\binom{q/p-1}{\lfloor k/p\rfloor}. \tag{4.5}$$

Proof. By the definition of R_k^c and the proof of Lemma 4.1 we have

$$R_k^c = \sum_{i_1=0}^{k}\sum_{i_2=0}^{k-i_1}\cdots\sum_{i_{c-2}=0}^{k-i_1-\cdots-i_{c-3}} R^2(k-\sum_{j=1}^{c-2}i_j),$$

where $R^2(k) = R_k^2$. From (3.2) and (3.3) it is easy to check that

$$R_k^2 - qS(k,b) \leq (q-p)\binom{q/p-1}{\lfloor k/p\rfloor}$$

for any $b \in \mathbf{F}_q$ when $p < q$. Therefore (4.5) follows since both the two numbers of terms appear in the two summations of R_k^c and S_k^c are $\binom{k+c-2}{c-2}$. □

Next we turn to giving a bound for R_k^c. Unfortunately, even though R_k^c can be written as a simple sum involving binomial coefficients, it seems nontrivial to evaluate it precisely. Using equation (4.1) and some combinatorial identities, we can easily obtain the following equality

$$R_k^c = -\sum_{j=0}^{\lfloor k/p\rfloor-1}(-1)^j\left[\binom{k+c-1-ip}{c-1}-\binom{k+c-1-ip-p}{c-1}\right]\binom{q/p-1}{j}$$
$$+\binom{< k >_p + c-1}{c-1}\binom{q/p-1}{\lfloor k/p\rfloor}. \tag{4.6}$$

It has been known that the simpler sum

$$\sum_{j=0}^{n}(-1)^j\binom{2n-1-3i}{n-1}\binom{n}{j},$$

which is the coefficient of x^n in $(1+x+x^2)^n$, has no closed form. That means it cannot be expressed as a fixed number of hypergeometric terms. For more details we refer to ([4], p. 160). This fact indicates that R_k^c also has no closed form. Thus, in the next lemma we just give a bound for R_k^c just using some elementary combinatorial arguments.

In Section 2 we have defined the unimodality of a sequence. A stronger property than unimodality is logarithmic concavity. First recall that a function f on the real line is concave if whenever $x<y$ we have $f((x+y)/2)\geq(f(x)+f(y))/2$. Similarly, a sequence $a_0,a_1\cdots,a_n$ of positive numbers is **log concave** if $\log a_i$ is a concave function of i which is to say that $(\log a_{i-1}+\log a_{i+1})/2\leq\log a_i$. Thus a sequence is log concave if $a_{i-1}a_{i+1}\leq a_i^2$. Using the properties of logarithmic concavity we have the following lemma.

Lemma 4.4.

$$R_k^c\leq p\cdot\max_{0\leq j\leq k}\binom{k+c-2-j}{c-2}\binom{q/p-1}{\lfloor j/p\rfloor}. \tag{4.7}$$

Proof. It is easy to check that both the two sequences $\binom{k+c-2-j}{c-2}$ and $\binom{q/p-1}{\lfloor j/p\rfloor}$ are log concave on j. Thus the sequence $a_j=\binom{k+c-2-j}{c-2}\binom{q/p-1}{\lfloor j/p\rfloor}$ is also log concave on j by the definition of logarithmic concavity . Since a log concave sequence must be unimodal, $\{a_j\}$ is unimodal on j. Then we have

$$\begin{aligned}R_k^c &= -\sum_{j=0}^{k}(-1)^{\lfloor j/p\rfloor}a_j\\ &= -\sum_{i=0}^{\lfloor k/p\rfloor}(-1)^i a_{ip}-\cdots-\sum_{i=0}^{\lfloor k/p\rfloor}(-1)^i a_{ip+<k>_p}\cdots-\sum_{i=0}^{\lfloor k/p\rfloor-1}(-1)^i a_{ip+p-1}.\end{aligned}$$

Thus (4.7) follows from the following simple inequality

$$\sum_{i=0}^{k}(-1)^i a_i\leq\max_{0\leq i\leq k}a_i,$$

and the proof is complete. □

Proof of Theorem 1.1 When $q>p$ we rewrite (4.3) to be

$$N(k,b,D)=\frac{1}{q}\binom{q-c}{k}+\frac{1}{q}(-1)^k(R_k^c-qM_k^c).$$

Applying (4.5) we obtain

$$\left|N(k,b,D)-\frac{1}{q}\binom{q-c}{k}\right|\leq\frac{q-p}{q}\binom{k+c-2}{c-2}\binom{q/p-2}{\lfloor k/p\rfloor}. \tag{4.8}$$

If $a_1=0$, and $b,a_2,\cdots,a_c$ are linear independent over $\mathbf{F}_p$, then $S_k^c=0$ for any k. Thus from (4.4) and Lemma 4.4 we have the improved bound

$$\left|N(k,b,D)-\frac{1}{q}\binom{q-c}{k}\right|\leq\frac{p}{q}\max_{0\leq j\leq k}\binom{k+c-2-j}{c-2}\binom{q/p-1}{\lfloor j/p\rfloor}. \tag{4.9}$$

Thus we only need to verify the case $q = p$. When $q = p$, from Lemma 4.1 we have

$$R_k^c = -\sum_{j=0}^{k} \binom{k+c-2-j}{c-2} = -\binom{k+c-1}{c-1}.$$

And $S(k, b)$ equals 0 or -1 by its definition given in Lemma 3.3. Thus from (4.3) we deduce that

$$N(k, b, D) = \frac{\binom{p-c}{k} - (-1)^k \binom{k+c-1}{k}}{p} + (-1)^k M_k^c \tag{4.10}$$

with $0 \le M_k^c \le \binom{k+c-2}{k}$. Thus

$$\left| N(k, b, D) - \frac{1}{q}\binom{q-c}{k} + \frac{(-1)^k}{q}\binom{k+c-1}{c-1} \right| \le \binom{k+c-2}{c-2}.$$

Note that $c = q - n$ and the proof is complete.

Example 4.5. *Choose $p = 2, q = 128, c = 4$ and $k = 5$. Then $R_k^c = -6840$. Let ω be a primitive element in $\mathbf{F}_{128}$. Let $D = F_{128}\backslash\{0, \omega, \omega^2, \omega^3\}$ and $b = 1$. Since $1, \omega, \omega^2, \omega^3$ are linear independent, (4.4) gives that there are $N = 1759038$ solutions of the equation (1.1) compared with the average number $\frac{1}{q}\binom{q-c}{k} \approx 1758985$.*

Remark. If one obtains better bounds for S_k^c, then we can improve the bound given by (4.8). However, it is much more complicated to evaluate S_k^c than R_k^c. Let

$$I = \{[i_1, i_2, \cdots, i_{c-2}], 0 \le i_t \le k - \sum_{j=1}^{t-1} i_j, 1 \le t \le c-2 : \quad b - \sum_{j=1}^{c-2} i_j a_{c+1-j} \in \mathbf{F}_p\}.$$

Simple counting shows that $0 \le |I| \le \binom{k+c-2}{c-2}$. In the proof of (4.8) we use the upper bound $|I| \le \binom{k+c-2}{c-2}$ and in the proof of (4.4) it is the special case $|I| = 0$. We can improve the above bound if we know more information about the cardinality of I, which is determined by the set $b, a_2, \cdots, a_c$. For example, if we know more about the rank of the set $\{b, a_2, \cdots, a_c\}$, then we can improve the bound given by (4.8). The details are omitted.

5. Applications to Reed-Solomon Codes

Let $D = \{x_1, \cdots, x_n\} \subset \mathbf{F}_q$ be a subset of cardinality $|D| = n > 0$. For $1 \le k \le n$, the Reed-Solomon code $D_{n,k}$ has the codewords of the form

$$(f(x_1), \cdots, f(x_n)) \in \mathbf{F}_q^n,$$

where f runs over all polynomials in $\mathbf{F}_q[x]$ of degree at most $k-1$. The minimum distance of the Reed-Solomon code is $n - k + 1$ because a non-zero polynomial of degree at most $k-1$ has at most $k-1$ zeroes. For $u = (u_1, u_2, \cdots, u_n) \in \mathbf{F}_q^n$, we can associate a unique polynomial $u(x) \in \mathbf{F}_q[x]$ of degree at most $n-1$ such that

$$u(x_i) = u_i,$$

for all $1 \le i \le n$. The polynomial $u(x)$ can be computed quickly by solving the above linear system. Explicitly, the polynomial $u(x)$ is given by the Lagrange interpolation formula

$$u(x) = \sum_{i=1}^{n} u_i \frac{\prod_{j \ne i}(x - x_j)}{\prod_{j \ne i}(x_i - x_j)}.$$

Define $d(u)$ to be the degree of the associated polynomial $u(x)$ of u. It is easy to see that u is a codeword if and only if $d(u) \leq k-1$.

For a given $u \in \mathbf{F}_q^n$, define

$$d(u, D_{n,k}) := \min_{v \in D_{n,k}} d(u, v).$$

The maximum likelihood decoding of u is to find a codeword $v \in D_{n,k}$ such that $d(u, v) = d(u, D_{n,k})$. Thus, computing $d(u, D_{n,k})$ is essentially the decision version for the maximum likelihood decoding problem, which is **NP**-complete for general subset $D \subset \mathbf{F}_q$. For standard Reed-Solomon code with $D = \mathbf{F}_q^*$ or $\mathbf{F}_q$, the complexity of the maximum likelihood decoding is unknown to be **NP**-complete. This is an important open problem. It has been shown by Cheng-Wan [2, 3] to be at least as hard as the discrete logarithm problem.

When $d(u) \leq k-1$, then u is a codeword and thus $d(u, D_{n,k}) = 0$. We shall assume that $k \leq d(u) \leq n-1$. The following simple result gives an elementary bound for $d(u, D_{n,k})$.

Theorem 5.1. *Let $u \in \mathbf{F}_q^n$ be a word such that $k \leq d(u) \leq n-1$. Then,*

$$n - k \geq d(u, D_{n,k}) \geq n - d(u).$$

Proof. Let $v = (v(x_1), \cdots, v(x_n))$ be a codeword of $D_{n,k}$, where $v(x)$ is a polynomial in $\mathbf{F}_q[x]$ of degree at most $k-1$. Then,

$$d(u, v) = n - N_D(u(x) - v(x)),$$

where $N_D(u(x) - v(x))$ denotes the number of zeros of the polynomial $u(x) - v(x)$ in D. Thus,

$$d(u, D_{n,k}) = n - \max_{v \in D_{n,k}} N_D(u(x) - v(x).$$

Now $u(x) - v(x)$ is a polynomial of degree equal to $d(u)$. We deduce that

$$N_D(u(x) - v(x)) \leq d(u).$$

It follows that

$$d(u, D_{n,k}) \geq n - d(u).$$

The lower bound is proved. To prove the upper bound, we choose a subset $\{x_1, \cdots, x_k\}$ in D and let $g(x) = (x - x_1) \cdots (x - x_k)$. Write

$$u(x) = g(x)h(x) + v(x),$$

where $v(x) \in \mathbf{F}_q[x]$ has degree at most $k-1$. Then, clearly, $N_D(u(x) - v(x)) \geq k$. Thus

$$d(u, D_{n,k}) \leq n - k.$$

The theorem is proved.

We call u to be a deep hole if $d(u, D_{n,k}) = n - k$, that is, the upper bound in the equality holds. When $d(u) = k$, the upper bound agrees with the lower bound and thus u must be a deep hole. This gives $(q-1)q^k$ deep holes. For a general Reed-Solomon code $D_{n,k}$, it is already difficult to determine if a given word u is a deep hole. In the special case that $d(u) = k+1$, the deep hole problem is equivalent to the subset sum problem over $\mathbf{F}_q$ which is **NP**-complete if $p > 2$.

For the standard Reed-Solomon code, that is, $D = \mathbf{F}_q^*$ and thus $n = q-1$, there is the following interesting conjecture of Cheng-Murray [1].

Conjecture Let $q = p$. For the standard Reed-Solomon code with $D = \mathbf{F}_p^*$, the set $\{u \in \mathbf{F}_p^n \big| d(u) = k\}$ gives the set of all deep holes.

Using the Weil bound, Cheng and Murray proved that their conjecture is true if p is sufficiently large compared to k.

The deep hole problem is to determine when the upper bound in the above theorem agrees with $d(u, D_{n,k})$. We now examine when the lower bound $n - d(u)$ agrees with $d(u, D_{n,k})$. It turns out that the lower bound agrees with $d(u, D_{n,k})$ much more often. We call u **ordinary** if $d(u, D_{k,n}) = n - d(u)$. A basic problem is then to determine for a given word u, when u is ordinary.

Without loss of generality, we can assume that $u(x)$ is monic and $d(u) = k + m$, $0 \le m \le n - k$. Let

$$u(x) = x^{k+m} - b_1 x^{k+m-1} + \cdots + (-1)^m b_m x^k + \cdots + (-1)^{k+m} b_{k+m}$$

be a monic polynomial in $\mathbf{F}_q[x]$ of degree $k + m$. By definition, $d(u, D_{n,k}) = n - (k+m)$ if and only if there is a polynomial $v(x) \in \mathbf{F}_q[x]$ of degree at most $k-1$ such that

$$u(x) - v(x) = (x - x_1) \cdots (x - x_{k+m}),$$

with $x_i \in D$ being distinct. This is true if and only if the system

$$\begin{aligned} \sum_{i=1}^{k+m} X_i &= b_1, \\ \sum_{1 \le i_1 < i_2 \le k+m} X_{i_1} X_{i_2} &= b_2, \\ &\cdots, \\ \sum_{1 \le i_1 < i_2 < \cdots < i_m \le k+m} X_{i_1} \cdots X_{i_m} &= b_m. \end{aligned}$$

has distinct solutions $x_i \in D$. This explains our motivational problem in the introduction section.

When $d(u) = k$, then u is always a deep hole. The next non-trivial case is when $d(u) = k + 1$. Using the bound in Theorem 1.1, we obtain some positive results related to the deep hole problem in the case $d(u) = k + 1$ (i.e., the case $m = 1$) if $q - n$ is small. When $q - n \le 1$, by Corollary 2.7 we first have the following simple consequence.

Corollary 5.2. *Let $q \ge n \ge q - 1$ and $q > 5$. Let $d(u) = k + 1$ with $2 < k < q - 3$. Then u cannot be a deep hole.*

Proof. By the above discussion, u is not a deep hole if and only if the equation

$$x_1 + x_2 + \cdots + x_{k+1} = b$$

always has distinct solutions in D for any $b \in F_q$. Thus the result follows from Corollary 2.7. □

Remark. Similarly, using Theorem 1.1, a simple asymptotic argument implies that when $q - n$ is a constant, and $d(u) = k + 1$ with $2 < k < q - 3$, then u cannot be a deep hole for sufficient large q. Furthermore, for given q, n, asymptotic analysis can give sufficient conditions for k to ensure a degree-$k + 1$ word u not being a deep hole.

In the present paper, we studied the case $m = 1$ and explored some of the combinatorial aspects of the problem. In a future article, we plan to study the case $m > 1$ by combining the ideas of the present papers with algebraic-geometric techniques such as the Weil bound.

References

[1] Q. Cheng and E. Murray, *On deciding deep holes of Reed-Solomon codes*, TAMS 2007, to appear.
[2] Q. Cheng and D. Wan, *On the list and Bounded distance Decodibility of Reed-Solomon Codes*, FOCS (2004), 335-341.
[3] Q. Cheng and D. Wan, *On the list and bounded distance decodability of Reed-Solomon codes*, SIAM J. Comput. **37** (2007), no. 1, 195-209.
[4] M. Petkovsek, H. S. Wilf and D. Zeilberger, $A=B$, Wellesley, MA:A. K. Peters, 1996.

School of Mathematical Sciences, Peking University, Beijing, P.R. China
E-mail address: joe@math.pku.edu.cn

Department of Mathematics, University of California, Irvine, CA 92697-3875, USA
E-mail address: dwan@math.uci.edu